UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/470,492	05/22/2009	William B. Dolan	327447.01	3989

69316 7590 04/28/2017

MICROSOFT CORPORATION
ONE MICROSOFT WAY
REDMOND, WA 98052

EXAMINER
KOVACEK, DAVID M

ART UNIT	PAPER NUMBER
2659	

NOTIFICATION DATE	DELIVERY MODE
04/28/2017	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

usdocket@microsoft.com
chriochs@microsoft.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte WILLAIM B. DOLAN, CHRISTOPHER J. BROCKETT, and LUCRETIA H. VANDERWENDE

Appeal 2016-006437
Application 12/470,492
Technology Center 2600

Before THU A. DANG, STEPHEN C. SIU, and JAMES W. DEJMEK, *Administrative Patent Judges*.

DANG, *Administrative Patent Judge*.

DECISION ON APPEAL

I. STATEMENT OF THE CASE

Appellants appeal under 35 U.S.C. § 134(a) from the Examiner's Final Rejection of claims 1–16 and 21–24. Claims 17–20 have been canceled. We have jurisdiction under 35 U.S.C. § 6(b).

We affirm.

A. INVENTION

According to Appellants, the invention is directed to a "mining system," which includes "cull[ing] a structured training set from an unstructured resource" that "may be latently rich in repetitive content and

alternation-type content" (i.e., the unstructured resource includes many instances of text that differ in form but express similar semantic content), (¶ 8).

B. REPRESENTATIVE CLAIM

Claim 1 is exemplary:

1. A method performed by a processing device, the method comprising:

constructing natural language queries comprising natural language query terms expressed in a natural language;

presenting the natural language queries to a web search engine, the web search engine configured to perform web crawling operations to maintain an index of web documents and use the index to identify matching web documents that match the natural language query terms of the natural language queries;

receiving web search result sets from the web search engine, the web search result sets providing web search result items identified by the web search engine from the matching web documents that match the natural language query terms of the natural language queries, wherein the web search result items are also expressed in the natural language; and

performing processing on the web search result sets to produce a training set, the training set identifying pairs of the web search result items within the web search result sets, wherein an individual pair in the training set includes:

a first web search result item received from the web search engine comprising first multiple words that are expressed in the natural language, and

a second web search result item received from the web search engine that comprises second multiple words that are also expressed in the natural language,

the training set providing a basis by which an electrical training system can learn a statistical translation model.

C. REJECTIONS

The prior art relied upon by the Examiner in rejecting the claims on appeal is:

Claims 1–12, 14, 15, and 21–24 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Brockett (US 2006/0106595 A1; pub. May 18, 2006), Niu (US 2009/0119090 A1; pub. May 7, 2009), and Decary (US 2007/0027672 A1; pub. Feb. 1, 2007).

Claim 13 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Brockett, Niu, Decary, and Stefik (US 2008/0027707 A1; pub. Jan. 31, 2008).

Claim 16 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Brockett, Niu, Decary, and Marcu (US 8,296,127 B2; pub. Oct. 23, 2012).

II. ISSUE

The principal issue before us is whether the Examiner erred in finding that the combination of Brockett, Niu, and Decary teaches or suggests "presenting the natural language queries to a web search engine" configured "to perform web crawling operations" to "identify matching web documents that match the natural language query terms" of the natural language queries. (Claim 1).

III. FINDINGS OF FACT

The following Findings of Fact (FF) are shown by a preponderance of the evidence.

Brockett

1. Brockett discloses a context model used to compare potential applicable paraphrase alternation patterns to the textual input. (Abstract). Brockett acknowledges that the recognition and generation of paraphrases is a key problem for applications of Natural Language Processing (NLP) system (¶ 3), and thus, provides a question answering system which uses paraphrase generation capability, for example, to produce multiple forms of a query to find matching results. (¶ 161).

Decary

2. Decary discloses using natural language processing to extract desired information from the Web page. (Abstract). Decary acknowledges that search engines that index millions of Web pages based on keywords are known ("Background of the Invention") that further accept natural languages queries, and then analyze the queries to extract the keywords for which the user is possibly looking. (¶ 25). Thus, Decary discloses using a lexical analysis (i.e., natural language processing), to find possible formal names on a given Web page and searching the given Web page using pattern matching techniques for formal names not found by the lexical analysis. (¶¶ 49–50). In Decary, a software robot "Crawler" is provided that "crawls" the Web, visiting and traversing the Web site with the goal of identifying and retrieving pages with relevant and interesting information (¶ 65), while a component is provided that performs data extraction on the pages retrieved by the Crawler based on Natural Language Processing techniques, wherein rules are used to identify and extract the relevant and interesting pieces of information. (¶ 66). Special rules are used to identify pages that contain organization information or relevant people information. (¶ 68).

IV. ANALYSIS

In reaching this decision, we consider all evidence presented and all arguments actually made by Appellants. We do not consider arguments that Appellants could have made but chose not to make in the Briefs, and we deem any such arguments waived. 37 C.F.R. § 41.37(c)(l)(iv).

Appellants contend "it is hard to see why Decary's description of a web search engine would lend one of ordinary skill to replace parallel corpora with search results from a web search engine." (App. Br. 15–16). In particular, Appellants contend "Decary's teachings that 'Web queries cannot become very specific' and 'tend to return thousands of results of which only a few may be relevant'" and thus are an "inadequate replacement for parallel corpora for the purpose of training a statistical translation model." (*Id.*, citing Decary at 25).

We have considered all of Appellants' arguments and evidence presented. However, we disagree with Appellants' contentions regarding the Examiner's rejections of the claims. Instead, we agree with the Examiner's findings, and find no error with the Examiner's conclusion that the claims would have been obvious over the combined teachings.

As an initial matter, as the Examiner points out, Appellants' argument regarding paragraph 25 of Decary "ignores the complete teachings of Decary," because "though [Appellants cite] the disclosure of Decary at {25}, such disclosure is included in Decary as a recitation of problems that exist without the conventional prior art," which "Decary seeks explicitly to improve upon." (Ans. 4).

We agree with the Examiner's finding "Decary provides further teachings, directed to the invention therein, that explicitly uses Web-based

data, extracted from the Web using a search engine [Crawler and Extractor] for training a grammar." (*Id.*; FF 2). In particular, "Decary itself specifically looks at particularly relevant web sources [traversing Web sites with the goal of identifying and retrieving pages with relevant and interesting information]," and "uses said sources to provide data matching within the context of a structure to be filled [lexical analysis to find formal names, which correspond to a particular structure; syntactically and grammatically identifying elements of interest, wherein syntax and grammar are understood to carry linguistic structure pertinent to the language providing such rules]." (Ans. 5; FF 2). That is, contrary to Appellants' contention that Decary's teachings "tend to return thousands of results of which only a few may be relevant" (App. Br. 15, citing Decary at 25), Decary solves the problem cited by Appellants by using a web crawler to retrieve relevant and interesting documents. (Ans. 5, FF 2).

Here, we find no error with the Examiner's reliance on Brockett for teaching and suggesting the steps of "presenting natural language queries to retrieval module" and "receiving result sets from the retrieval module." (Final Rej. 4; FF 1). Even Appellants do not contest the Examiner's findings in Brockett.

Further, Decary discloses, using natural language processing to extract relevant/interesting information and applying pattern matching techniques, a method which comprises crawling the Web to identify and retrieve pages with relevant and interesting information and performing data extraction on the pages retrieved by the Crawler. (FF 2). Thus, as discussed above, we find no error with the Examiner's reliance on Decary for teaching and suggesting "a retrieval module . . . configured to perform web crawling

operations" to "identify matching we[b] documents that match the natural language query terms of the natural language queries." (Final Rej. 6–7; FF 2).

Accordingly, we find no error with the Examiner's finding that the combination of Brockett and Decary teaches or at least suggests "presenting the natural language queries to a web search engine" configured "to perform web crawling operations to maintain an index of web documents" to "identify matching web documents that match the natural language query terms" of the natural language queries. (Claim 1).

The Supreme Court has clearly stated the "combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results." *KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398, 416 (2007). That is, when considering obviousness of a combination of known elements, the operative question is thus "whether the improvement is more than the predictable use of prior art elements according to their established functions." *Id.* at 417. The skilled artisan is "a person of ordinary creativity, not an automaton." *Id.* at 421.

Although Appellants contend "it is hard to see why Decary's description of a web search engine would lend one of ordinary skill to replace parallel corpora with search results from a web search engine" (App. Br. 15), Appellants appear to view the combination from a different perspective than that of the Examiner. Here, Appellants have presented no evidence that combining Brockett's natural language queries for receiving result sets to Decary's web crawling operations to identify matching web documents that match natural language terms would have been "uniquely challenging or difficult for one of ordinary skill in the art." *Leapfrog*

Enters., Inc. v. Fisher-Price, Inc., 485 F.3d 1157, 1162 (Fed. Cir. 2007) (citing KSR, 550 U.S. at 420). Instead, we agree that Appellants' invention is simply a modification of familiar prior art teachings (as taught or suggested by the cited references) that would have realized a predictable result. *KSR* 550 U.S. at 421.

On this record, we are unconvinced of Examiner error in the rejection of independent claim 1 over Brockett and Decary.

As to claim 12, Appellants merely contend Decary does not render obvious "extracting the natural language query terms from a web-based encyclopedic reference source." (App. Br. 18, repeating the language of claim 12). However, as the Examiner points out, the Examiner is relying on the combination of Brockett and Niu, as well as Decary, to at least suggest the limitation. (Ans. 7–8). Furthermore, we agree with the Examiner that "Decary does explicitly state that data is extract from Web sources" and "that 'old books and journals' can be found on the Web publicly." *Id.* Thus, we agree with the Examiner that "one of ordinary skill in the art would recognize that the sources . . . would be included within the Web content used for extraction of relevant data" in Decary. *Id.* On this record, we are not persuaded that the Examiner erred in this finding.

Appellants do not provide separate arguments for the other pending claims, and thus, we summarily affirm the rejections of these claims. *See* 37 C.F.R. § 41.37(c)(1)(iv).

V. CONCLUSION AND DECISION

We affirm the Examiner's rejection of claims 1–16 and 21–24 under 35 U.S.C. § 103(a).

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a)(1)(iv).

AFFIRMED